A novel test to differentiate anosmic malingerers from actually anosmic patients

Jalal Mehdizade, M.D.,[1] Babak Saedi, M.D.,[1] Reza Fotouhi, M.D.,[2] and Amin Safavi, M.D.[2]

ABSTRACT

Background: *The available olfactory evaluation tests are mainly subjective methods requiring patients' collaboration. If, for any reason, the patients refuse to honestly report what they perceive, the test reliability will be questionable; this condition is potentially observable in malingering patients because of their financial or psychosocial incentives. In an olfactory discrimination test context, this study was aimed to design a test capable of distinguishing malingerer from actually anosmic or severely hyposmic patients.*

Methods: *The pilot experiment of our methodology study determined five substances (coffee, lemon, rosewater, thyme, and garlic) as qualified odors of a 20-item odor discrimination test and set its normal reference value at 15. Through two simulations, 70 normosmic participants emulated actual anosmia and also malingering. The outcome results were used to measure test reliability factors.*

Results: *During the malingering simulation, only seven participants were capable of keeping their scores at the test chance level with enough randomness in their sequences of answers while the actual anosmia simulation revealed that 39 had scores at the test chance level. Accordingly, the Tehran University Odor Discrimination Test (TUODT) was measured to have 90% sensitivity, 55.71% specificity, 67.02% positive predictive value, and 84.78%negative predictive value.*

Conclusion: *The TUODT is a relatively efficient method to identify anosmia malingerers.*

(Am J Rhinol Allergy 26, 485–488, 2012; doi: 10.2500/ajra.2012.26.3812)

The conventional olfactory evaluation tests are mainly subjective methods assessing olfactory threshold, odor identification, smell memory, and odor discrimination. All of these tests need patient's collaboration because it is the patient who has to perceive, nominate, recall, and distinguish between the presented odors.[1] Accordingly, the reliability of these tests is directly dependent on the patient's tendency to honestly report his/her perceptions during the test or to intentionally skew the results. Especially in cases involving litigation, patients with acquired olfactory impairments are prone to feign anosmia because if they validate their anosmic condition due to iatrogenic interventions, head traumas, or occupational exposures, a huge payment of disability compensation in response to insurance commitments or court's decision would be guaranteed.

In the context of subjective studies, olfactory irritant odors,[2,3] malingerers' personal characteristics,[4]and test response characteristics[2,5,6]are used to distinguish malingerers from actually anosmic patients. Optimally, we could precisely evaluate one's olfactory function and localize the possible lesion if we had meticulous objective methods such as what we have for audition (auditory brainstem response) or vision (visual evoked potentials). In this category, trials have electrically or chemically stimulated the olfactory neural system[7–9]with the capability of detecting corresponding brain-evoked potentials, but, to date, none have been qualified for clinical use.

Therefore, the current evaluation methods are incapable of determining patients' honesty in their responses unless we use them as psychophysical tests. To design a test that is subjectively capable of distinguishing malingerers from actual anosmic patients, we conducted a methodology study based on the odor discrimination discipline and in a simulation setting we studied the possible reactions of malingerers and anosmic patients. The outcome results were used as test reliability factors in measuring test sensitivity, specificity, and also positive and negative predictive values.

From the Departments of [1]Otolaryngology and [2]Ear, Nose, and Throat, Ear, Nose, and Throat Research Center, Tehran University of Medical Sciences, Tehran, Iran
The authors have no conflicts of interest to declare pertaining to this article
Address correspondence and reprint requests to Babak Saedi, Otolaryngology Research Center, Imam Khomeini Medical Center, Bagherkhan Street, Chamran Highway, Tehran, Iran 141973141
E-mail address: saedi@tums.ac.ir

SUBJECTS AND METHODS

This study was aimed to design a test capable of distinguishing between actually anosmic patients and malingerers; therefore, we conducted a methodology study at a tertiary health care center (Imam Khomeini Hospital, an affiliate of Tehran University of Medical Sciences) through the years 2010 to 2011. The study had two distinct phases: a pilot experimentation and an odor discrimination test and simulation series.

The Pilot Experiment and Calibration

The pilot experiment, as a small-scale preliminary study, was settled to determine the test materials and obtain normal reference values; therefore, a series of 15 normosmic people with no previous positive history of olfactory impairment or diseases altering olfaction (such as diabetes, head traumas,[10,11]sinonasal disorders,[3] brain degenerative diseases,[12] *etc.*) were selected from among a normal population.

Evaluation of Olfactory Thresholds. As the first step in the experiment, olfactory thresholds of the participants were evaluated through a three-odor series presentation; this series consisted of two distilled water bottles (odorless) and a 0.25% phenyl ethyl alcohol bottle (the least perceptible intensity of phenyl ethyl alcohol by normal people).[13] Participants with normal olfactory thresholds were eligible to enter the calibration step of the experiment as were all of the 15 pilot experiment entrants.

Test Calibration. Several potentially acceptable odorants, all arranged in the three-odor series format, were pooled and presented to the participants. Some were not familiar to them and some could confuse them because of their similarity; therefore, through omission of these odors, the following five odors were considered ideal for the test: coffee, lemon, rosewater, thyme, and garlic.

Study Materials

The five selected odors were all provided in unlabeled identical dark bottles in aqueous solutions at higher intensities than the normal human olfactory threshold (suprathreshold test). All bottles contained 10 mL of odorants materials. Twenty discrimination test items were made out of these five odors, each comprising of three odors; therefore, our 20-item test kit contained a total of 60 bottles. The purpose of bottle arrangement was to put two identical odors in the same item while the third was different; therefore, two series of the basic five

Table 1 **Contents of a 20-item test kit**

R, 1-6-2	L, 2-5-7	L, 3-8-2	R, 4-8-3	L, 5-9-4
L, 5-10-3	L,1-4-6	R, 2-7-3	L, 3-1-8	R,2-1-7
R, 5-6-1	L,4-9-1	L,2-5-10	R,5-1-10	R,3-8-5
R, 4-2-9	L,3-6-1	R, 2-7-4	R,5-4-10	L,4-9-3

The different odors in each series are underlined.
L = left; R = right.

odors were needed for each item to select the identical odors (effectively, three basic odor series for 5 items, 12 series for a total of 20-item test kit). Accordingly, for a single item, the two series of basic five odor bottles were numbered from 1 to 10; therefore, bottle numbers for the coffee odor were 1 and 6; lemon, 2 and 7; rosewater, 3 and 8; thyme, 4 and 9; and garlic, 5 and 10. On the other hand, each test item was presented to a single nostril; therefore, 10 items were randomly entitled as L (left) and the rest as R (right). Table 1 shows the 20-item test kit.

The participants had to determine the different odor in each series by numbers according to bottle presentation sequence; *e.g.*, in the three-odor series of thyme–thyme–coffee (the L, 4-9-1 item), the right answer would be number 3. Each odor was exposed at a 2-cm distance from the nostril and only for 3 seconds. A 30-second interitem interval was also included to prevent olfactory exhaustion and adaptation. To avoid bottle content degeneration, the odors kits were renewed each month.

Normal Reference Values

At optimum conditions, a participant could earn a total of 20 scores from the odor discrimination test. The 15 pilot experiment participants again took the new test in its new format, and scored at least 15; therefore, the lower normal limit for a discrimination test score was set at 15. On the other hand, due to the 0.33 chance-level of the three-choice test items, scores of six and seven were considered to be at the test chance-level.

Odor Discrimination Tests and Simulations

The second phase of the study consisted of three components: the odor discrimination test, the malingering simulation, and the anosmia simulation.

1. Odor Discrimination Test, the Study Subject Selection. One hundred candidates (none of whom attended the pilot study), aged 18–45 years (to minimize the age-related decline in olfactory function) and with higher education (university students or graduates; possibly having enough intelligence to decipher the test discipline) were recruited for this section. They were all healthy normosmic individuals. These candidates undertook the 20-item odor discrimination test; those who scored at least 15 were considered to have a normal olfactory function and were eligible to enter the simulations. Finally, 70 eligible participants entered the study.
2. Malingering Simulation. Knowing the fact that the test had a 0.33 chance level and based on a 20-score scale, participants were asked to intentionally limit their scores to 6 or 7 (the test chance level). They were also notified to avoid a logical sequence in their correct answers, which would unveil their intention to skew the results; in other words, they were asked to randomly answer correctly to one-third of the items. This was how the intelligent enough malingerers would act and would not be diagnosed as malingerers (false negatives) and the rest would be diagnosed as malingerers (true positives). Accordingly, this simulation was to determine the participants' capability of malingering successfully.

Table 2 **Test characteristics and reliability measurements**

Variables	Definition	Values
Components		
No. of MS scores other than 6 and 7	TP	63
No. of AS scores other than 6 and 7	FP	31
No. of MS scores of 6 and 7	FN	7
No. of AS scores of 6 and 7	TN	39
Equations		
Sensitivity	TP/(TP + FN)	90.00%
Specificity	TN/(TN + FP)	55.71%
Positive predictive value	TP/(TP + FP)	67.02%
Negative predictive value	TN/(TN + FN)	84.78%

MS = malingering simulation; AS = anosmia simulation; FN = false negative; TP = true positive; FP = false positive; TN = true negative.

3. Anosmia Simulation.

No odor in this section was presented to the participants while they were asked to randomly select the choices one, two or three for each of the 20 items. This would simulate the way that anosmic patients would select the provided choices while they did not really have the ability to smell. The results of this simulation were supposed to be at the test chance level (true negatives) whereas some would obtain other scores (false positives).

Other Variables

All of the results throughout the study were recorded into provided code sheets and, accordingly, the data were analyzed to determine how effectively the test could determine the malingering patients. Therfore, the test sensitivity, specificity, positive predictive value, and negative predictive value were calculated. Patients were categorized into three age groups according to their age distribution. Moreover, the olfactory scores were compared in age groups, sex groups, and nostril laterality.

Ethical Approval

The protocol of this study was approved by the Institutional Review Board of the Tehran University of Medical Sciences. All aspects of the study were conducted according to the Declaration of Helsinki. All of the entrants were aware of the investigational nature of the study and that they would not undergo any medical or surgical intervention and that the confidentiality of their test results would be maintained. They all agreed to participate by signing our printed informed consent.

Statistical Methods

Table 2 depicts how the test sensitivity, specificity, and positive and negative predictive values were measured. The sensitivity of the Tehran University Odor Discrimination Test (TUODT) means the percentage at which the test could correctly identify the malingerers; and its specificity denotes the percentage at which the test could identify the actually anosmic patients. Furthermore, the positive predictive value describes the probability at which a patient could really be malingering and the test correctly recognized him/her as a malingerer; whereas the negative predictive value represents the probability at which the test could correctly identify a real case of anosmia as anosmic.

On the other hand, the Kruskai-Wallis one-way analysis of variance was used to statistically evaluate the correlation between age groups and mean odor discrimination scores while the Mann-Whitney *U* test was used to analyze the possible correlation between odor discrimi-

Table 3 **Mean odor discrimination test scores in patient groups**

Characteristics	Mean Odor Discrimination Test Score
Gender	
Male	17.31
Female	16.88
Total	17.15
Age	
18–25 yr	17.5
26–33 yr	17.09
Over 33 yr	16.75
Nostril	
Right	8.27
Left	8.87

nation scores and patients' gender and nostril laterality. The data were analyzed using SPSS Version 11.5 (SPSS, Inc., Chicago, IL). Values of $p \leq 0.05$ were considered significant.

RESULTS

A total of 70 participants (25 male and 45 female subjects) with a mean age of 29.1 ± 3.5 years (range, 18–45 years) entered the simulations. All of the participants were selected from among healthy adults with higher education (medical university students or graduates) who had obtained scores of ≥15 through the odor discrimination test.

Odor Discrimination Test Scores

Table 3 summarizes mean odor discrimination scores according to patients' gender, age, and nostril laterality.

The mean odor discrimination test scores showed no significant correlation with age (Kruskal-Wallis test, $p = 0.368$), participants' gender, or nostril laterality (Mann-Whitney U test, $p = 0.317$ for both variables).

Scores of Malingering and Anosmia Simulations

As outlined in Table 4, of the 10 individuals who had scored 6 or 7 in malingering simulation, two had relatively logical orders in their answer sequences and one had a completely logical order to follow; therefore, the total number of the individuals who were definitely capable of malingering decreased to 7 (10% of total) who could all restrict their scores to 6 and 7 without any obvious logical pattern. Through the random selection of choices in the real anosmia simulation, 31 (44%) individuals had scores higher than expected at the study chance level.

Test Characteristics

According to Table 2, the 90% sensitivity of TUODT means that it was capable of identifying up to 90% of the malingering individuals, and its 55.71% specificity showed the capability of the test to identify actually anosmic patients. On the other hand, the 67.02% positive predictive value of the test represented the probability of malingering when the test recognized it as so and its 84.78% negative predictive value described the probability of actual anosmia when the test diagnosed a patient as being so.

DISCUSSION

It is an important issue to reliably identify a patient's problem for its further treatment. Occasionally, the test results may be in contrast with patients' gains and external incentives, so they may try to exaggerate their symptoms or underreport the questioned factors in a test,[4] which clearly alters the test reliability. In these cases, tests that can possibly exempt patients' intervention and objectively identify

Table 4 **Scores distribution in tests**

Test	Scores >7	Scores of 6 or 7	Scores <6
Odor discrimination test	70 (100%)	0	0
Anosmia malingering simulation scores	54 (77%)	10 (14%)	6 (9%)
Real anosmia simulation scores	15 (21%)	39 (56%)	16 (23%)

the problem are invaluable. Because the current olfactory evaluation tests are based on subjective methods and no available objective test is yet qualified for olfactory evaluation, we used a subjective test to psychophysically detect patients' willingness to intentionally interfere with the test results.

Odor familiarity is the most important factor on which odor-identification tests are dependent,[14] although a well-known odor in one culture may not be quite as well-known in another, making it a cultural phenomenon.[15] The discriminative nature of our study makes it a potential option for intercultural use because odor discrimination tests do not need to know the name of the odors used, and what the patients have to do is simply discriminate the one different odor among the other identical ones.

The odor discrimination test was used as a tool to qualify the participants who entered simulations, and the test could efficiently identify the required normosmic people. We also analyzed the participants' characteristics with the same test. In the literature, women are reported to have a better olfactory function than men[16–19]and the left nostril is reported to outperform the right[16]; female subjects and the left nostrils had higher scores than men and the right nostrils in our study, but these findings did not reach the level of statistical significance. The same was true for olfactory scores in different age groups although there are reports of smell changes with aging.[1,19–23] The reason for the aforementioned findings might be that in the first place, the patients were assessed with odors previously qualified to be at a high contrast with each other in the pilot study and second, the patients were selected among individuals who were too young to have an age-related olfactory deficit.

Our study showed a 90% sensitivity to identify the malingerers; although it is at an acceptable level, the test may lose its validity if the participants have the ability to decipher its discipline. This is because the test had a predetermined number of items (it is a 20-item test), and what the participant needed to know for successful malingering was the accurate number of test items to calculate the supposed scores at the test chance level. Accordingly, the test length seems to be a key factor causing bias[5]; for this purpose, blinding the number of test items by using test series with different numbers of items can be a solution that can definitely increase the test validity. On the other hand, with a fixed number of test items, analyzing the randomness in the sequence of responses[6] can signify the possible intention of the patients to give wrong answers because we could identify three of our participants who had logical orders in their response sequences. The 55.71% specificity of the test describes the relatively low potential of the test to determine anosmic patients; we did not expect a higher specificity for this test because this test was sensitive to intentionally manipulated data.

Moreover, if we evaluate the logical pattern of false answers, the malingered participant can better be compared with crude answers. Therefore, it seems that crude answers should not be the only source of data uses in final analysis of such tests, but logical patterns can be revealed by evaluating the wrong answers and be very useful data in test interpretations. Additionally, if the special software can find the relationship of wrong answers, it can be useful information.

Although the future objective olfactory evaluation methods would be the gold standard for identifying malingering patients, the TUODT as a subjective psychophysical test could distinguish malingerers from actually anosmic patients with an acceptable sensitivity.

CONCLUSION

As a psychophysical olfactory evaluation test, the TUODT can serve as a relatively competent method to diagnose malingerers of anosmia. Additional larger-scale surveys are needed to confirm its efficacy.

REFERENCES

1. Gaines AD. Anosmia and hyposmia. Allergy Asthma Proc 31:185–189, 2010.
2. Kurtz DB, White TL, Hornung DE, and Belknap E. What a tangled web we weave: Discriminating between malingering and anosmia. Chem Senses 24:697–700, 1999.
3. Kim DW, Kim JY, and Jeon SY. The status of the olfactory cleft may predict postoperative olfactory function in chronic rhinosinusitis with nasal polyposis. Am J Rhinol Allergy 25:90–94, 2011.
4. Doty RL, and Crastnopol B. Correlates of chemosensory malingering. Laryngoscope 120:707–711, 2010.
5. Doty RL. Olfactory dysfunction and its measurement in the clinic and workplace. Int Arch Occup Environ Health 79:268–282, 2006.
6. Linschoten MR, and Harvey LO Jr. Detecting malingerers by means of response-sequence analysis. Percept Psychophys 66:1190–1201, 2004.
7. Ishimaru T Shimada T, Sakumoto M, et al. Olfactory evoked potential produced by electrical stimulation of the human olfactory mucosa. Chem Senses 22:77–81, 1997.
8. Lago P, Broich G, Cecchini AP, et al. Trigeminal evoked potentials in man: A new olfactory stimulation device. Funct Neurol 13:135–140, 1998.
9. Wetter S, and Murphy C. A paradigm for measuring the olfactory event-related potential in the clinic. Int J Psychophysiol 49:57–65, 2003.
10. Reiter ER, DiNardo LJ, and Costanzo RM. Effects of head injury on olfaction and taste. Otolaryngol Clin North Am 37:1167–1184, 2004.
11. Yamagishi M, Okazoe R, and Ishizuka Y. Olfactory mucosa of patients with olfactory disturbance following head trauma. Ann Otol Rhinol Laryngol 103:279–284, 1994.
12. Fleiner F, Dahlslett SB, Schmidt F, et al. Olfactory and gustatory function in patients with multiple sclerosis. Am J Rhinol Allergy 24:e93–e97, 2010.
13. Lotsch J, Lange C, and Hummel T. A simple and reliable method for clinical assessment of odor thresholds. Chem Senses 29:311–317, 2004.
14. Shu CH, and Yuan BC. Assessment of odor identification function in Asia using a modified "Sniffin' Stick" odor identification test. Eur Arch Otorhinolaryngol 265:787–790, 2008.
15. Saito S, Ayabe-Kanamura S, Takashima Y, et al. Development of a smell identification test using a novel stick-type odor presentation kit. Chem Senses 31:379–391, 2006.
16. Cardesin A, Alobid I, Benítez P, et al. Barcelona Smell Test-24 (BAST-24): Validation and smell characteristics in the healthy Spanish population. Rhinology 44:83–89, 2006.
17. Doty RL, Applebaum S, Zusho H, and Settle RG. Sex differences in odor identification ability: A cross-cultural analysis. Neuropsychologia 23:667–672, 1985.
18. Doty RL, Shaman P, and Dann M. Development of the University of Pennsylvania Smell Identification Test: A standardized microencapsulated test of olfactory function. Physiol Behav 32:489–502, 1984.
19. Kobal G, Hummel T, Sekinger B, et al. "Sniffin' sticks": Screening of olfactory performance. Rhinology 34:222–226, 1996.
20. Doty RL, Shaman P, Applebaum SL, et al. Smell identification ability: Changes with age. Science 226:1441–1443, 1984.
21. Hummel T, Sekinger B, Wolf SR, et al. "Sniffin' sticks": Olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. Chem Senses 22:39–52, 1997.
22. Paik SI, Lehman MN, Seiden AM, et al. Human olfactory biopsy. The influence of age and receptor distribution. Arch Otolaryngol Head Neck Surg 118:731–738, 1992.
23. Schiffman SS. Taste and smell losses in normal aging and disease. JAMA 278:1357–1362, 1997. □